

Generalized sampling reconstruction from Fourier measurements using compactly supported shearlets

Jackie Ma*

Technische Universität Berlin
Department of Mathematics
Straße des 17. Juni 136, 10623 Berlin

July 31, 2015

Abstract

In this paper we study the general reconstruction of a compactly supported function from its Fourier coefficients using compactly supported shearlet systems. We assume that only finitely many Fourier samples of the function are accessible and based on this finite collection of measurements an approximation is sought in a finite dimensional shearlet reconstruction space. We analyse this sampling and reconstruction process by a recently introduced method called *generalized sampling*. In particular by studying the *stable sampling rate* of generalized sampling we then show stable recovery of the signal is possible using an almost linear rate. Furthermore, we compare the result to the previously obtained rates for wavelets.

1 Introduction

A general task in sampling theory is the reconstruction of an object from finitely many measurements. Such a problem also appears in more applied fields such as signal processing and medical imaging, the latter being a main motivation of this paper with view to applications. A Hilbert space where such problems are often modeled in is the Hilbert space $L^2(\mathbb{R}^2)$ - the space of square integrable functions -, indeed, the reconstruction task in magnetic resonance imaging (MRI) is mostly modeled in $L^2(\mathbb{R}^2)$. In this paper we assume the object of interest can be modeled by a compactly supported function $f \in L^2(\mathbb{R}^2)$ and the acquired samples are simply the Fourier coefficients of this function. Of course, a reconstruction of such a function f can then be obtained by forming a truncated Fourier series using the sampled Fourier coefficients. However, physical constraints make it impossible to sample all Fourier coefficients which causes an error in the reconstruction and possibly leads to non satisfactory reconstructions. For example the Gibbs phenomenon is a typical artifact that is strongly visible in the approximation if not enough Fourier coefficients are used in a Fourier series. Sampling Fourier coefficients and forming a reconstruction by a Fourier series means that sampling and reconstructing the signal is both done by using the same system. While the way we acquire the data is indeed usually fixed by the acquisition device, for instance in MRI, the reconstruction system is not. Very often additional information about the object is known, e.g. the shape or the regularity. Therefore it is plausible to use this extra information in the reconstruction scheme by choosing a suitable reconstruction system that favors a representation of this object.

*ma@math.tu-berlin.de

1.1 Sampling and reconstruction model

As already mentioned at the beginning of the introduction, many reconstruction problems can be modeled in a Hilbert space \mathcal{H} with inner product $\langle \cdot, \cdot \rangle$. The samples are then assumed to be acquired with respect to a fixed sampling system $\{s_1, s_2, \dots\} \subset \mathcal{H}$ and the reconstruction is computed using another system $\{r_1, r_2, \dots\} \subset \mathcal{H}$ which can be arbitrarily chosen. The concrete reconstruction problem can then be formulated as follows: Given finitely many linear measurements $\langle f, s_1 \rangle, \dots, \langle f, s_M \rangle, M \in \mathbb{N}$ of a function $f \in \mathcal{H}$, find a reconstruction $f_N \in \text{span}\{r_1, \dots, r_N\}, N \in \mathbb{N}$ that is numerically stable and whose error converges to zero as N tends to infinity.

Reconstruction problems of this type have already been studied, e.g. by Unser and Aldroubi in [34] and later also by Eldar in [15, 16] for the case $N = M$ where the used method is called *consistent reconstruction*. We investigate this reconstruction problem by using *generalized sampling* (GS) which was introduced by Adcock et al. in a series of papers [3, 4, 6]. In fact, GS can be seen as an extension of consistent reconstruction and a comparison of both methods is presented in [6]. The key difference of GS to consistent reconstruction is to allow M and N to vary independently of each other. This additional flexibility can be used to overcome some of the issues that are addressed in [6]. GS further provides a quantity called the *stable sampling rate*, which represents the number of samples $M \in \mathbb{N}$ that are needed in order to obtain an approximation that is stable and convergent using the first $N \in \mathbb{N}$ reconstruction elements of the a priori chosen and ordered reconstruction system. We will introduce generalized sampling and in particular the stable sampling rate in sufficient detail in Section 2.

1.2 Sparsifying systems

The selection of a different system for reconstruction than the given sampling system can only be done faithfully if the structure of the underlying object is known. Indeed, it is well known that most natural images, in particular medical images, have a *sparse representation* with respect to a *wavelet basis* [12], meaning either many wavelet coefficients are zero or have a fast decay with respect to increasing scales. This sparse representation by wavelets is one of the reasons why they became very important in imaging problems, see e.g. [33].

However, over the years many other sparsifying systems similar to wavelets have arisen such as *curvelets* [9] and *shearlets* [24]. More importantly, these systems are known to outperform wavelets in higher dimensions in the sense that both, curvelets and shearlets provide a so-called *almost optimal sparse approximation rate* of *cartoon-like* images which are a standard model for natural and medical images, see [13, 9, 25], while wavelets provably do not reach this rate. For this reason, it is of utmost importance to study not only wavelets but in addition shearlets as a reconstruction system to approximate a compactly supported function based on its Fourier coefficients in the context of generalized sampling. Indeed, wavelets has been studied in [7, 5] see also Subsection 1.3.

Note that there exists a general notion of *parabolic molecules* introduced by Grohs et al. [19] for which shearlets and curvelets are special instances. A natural question is then, whether our results can be extended to such systems and we will comment on this in Section 7.

1.3 Contribution and related work

To enhance the quality of the Fourier inversion of the sampled data, great effort has been done in applications to improve the reconstruction process by changing the reconstruction basis or adapting the sampling process. For example in [21] the authors introduced *wavelet-encoding* which allows the user to directly sample wavelet coefficients instead of Fourier coefficients by adapting the acquisition device. However, as for now this method also has several shortcomings such as low signal-to-noise ratio [30]. On the other hand, the proposed method GS can be

understood as a post processing method that does not require to change the acquisition device since the sparsifying transform will be incorporated in the reconstruction process and not in the sampling process.

The use of GS to study the reconstruction problem from Fourier measurements has already been investigated for wavelets in 1D [7] and 2D [5]. The authors of the respective works showed, that for dyadic multiresolution analysis (MRA) wavelets the stable sampling rate is linear, meaning up to a constant one has to sample as many Fourier coefficients as one wants to reconstruct wavelet coefficients. Moreover, it was shown in [7] that a defiance of this rate leads to unstable approximations.

The main result of our paper shows that the stable sampling rate is almost linear for sampling Fourier coefficients and reconstructing shearlet coefficients. The main challenge compared to the previous results of compactly supported MRA wavelets is that shearlets neither obey a classical scaling equation which enables an expansion of translations of a single function at highest scale nor form an orthonormal bases. However, both properties are heavily used in the proofs for the wavelet case. Therefore the proof requires a different strategy and is build upon good localization properties in frequency of the reconstruction system. Moreover, to the best of the authors knowledge there are so far no results for a stable sampling rate for redundant reconstruction systems that merely form a frame. In fact, the redundancy or instability measured by the lower frame bound of the reconstruction system now becomes apparent and effects the stable sampling rate directly. Such a penalty does not occur when using orthonormal bases or Riesz bases. Therefore the rate that we obtain for shearlets is worse than the rate for wavelets.

Since the stable sampling rate is worse for shearlets than for wavelets we justify the theoretical gain of using shearlets in combination of GS by showing that the error of the GS shearlet reconstruction converges asymptotically faster to zero than the error of the GS wavelet reconstruction under some assumptions if both reconstructions are computed from the same number of measurements.

As for practical applications such as MRI, wavelets are often used as a reconstructions system and it has been shown that these often lead to superior results in terms of lower sampling rates and reduced artifacts [28]. We will provide some numerical tests with MR data and deterministic subsampling patterns in the numerics section using shearlets and compare these to reconstruction methods based on wavelets and Fourier inversion to show the strength of using shearlets as a reconstruction system for the recovery problem from Fourier measurements.

1.4 Notation

In GS the sampling system and the reconstruction system are allowed to be almost completely arbitrary, in fact, assuming the frame property is enough. Recall that a countable sequence $(f_i)_{i \in I}$ is called a *frame* ([11]) for a Hilbert space \mathcal{H} with inner product $\langle \cdot, \cdot \rangle$ if there exist $0 < A \leq B < \infty$ such that

$$A\|f\|^2 \leq \sum_{i \in I} |\langle f, f_i \rangle|^2 \leq B\|f\|^2, \quad \forall f \in \mathcal{H}.$$

For most parts we will assume \mathcal{H} to be the space of square integrable functions $L^2(\mathbb{R}^2)$ equipped with the standard inner product $\langle f, g \rangle := \int f \overline{g} d\mu$ and induced norm $\|f\| := \sqrt{\langle f, f \rangle}$ for $f, g \in L^2(\mathbb{R}^2)$.

The Fourier transform of a function $f \in L^1(\mathbb{R}^2)$ is denoted by \widehat{f} and we use the definition

$$\widehat{f}(\xi) = \int_{\mathbb{R}^2} f(x) e^{-2\pi i \langle x, \xi \rangle} dx, \quad \xi \in \mathbb{R}^2,$$

and with a standard approximation argument we extend this definition to functions in $L^2(\mathbb{R}^2)$.

We will also often write " \lesssim " to indicate an inequality to hold up to some constant that is irrelevant or independent of all important parameters. Similarly " \asymp " should denote equality up to some constant.

1.5 Outline

We continue with a short review of generalized sampling in Section 2. Then we introduce shearlets and recall the concept of sparse approximation of cartoon-like functions in sufficient detail in Section 3. Section 4 contains the precise definitions of the sampling and reconstruction spaces that are used for the main theorem that is formulated in Subsection 4.3. All proofs are then presented in Section 5. Finally, in Section 6 we aim to numerically confirm the methodology and effectiveness of using shearlets as a reconstruction system.

2 Generalized sampling

Generalized sampling was introduced by Adcock et al. in [3, 4, 6] as a reconstruction method for almost arbitrary sampling and reconstruction systems. In fact, similar to *consistent reconstruction* [34, 15, 16] these systems are assumed to form a frame. As already mentioned in the introduction, the key difference of generalized sampling and consistent reconstruction is to allow the number of measurements to vary independently of the number of reconstruction elements. This flexibility enables generalized sampling to overcome some of the barriers of consistent reconstruction [3].

2.1 GS reconstruction method

Given a *sampling system* $\{s_1, s_2, \dots\} \subset \mathcal{H}$ we define the sampling space $\mathcal{S} \subset \mathcal{H}$ as the closure of its span, i.e.

$$\mathcal{S} = \overline{\text{span}}\{s_k : k \in \mathbb{N}\}.$$

The finite dimensional version of this sampling space is denoted by

$$\mathcal{S}_M = \text{span}\{s_1, \dots, s_M\}, \quad M \in \mathbb{N}.$$

The sampling vectors $\{s_k : k \in \mathbb{N}\}$ are used to model the measurements, in fact, we assume that the samples are given as linear measurements of the form

$$m_f(k) = \langle f, s_k \rangle, \quad k \in \mathbb{N}. \quad (2.1)$$

Analogously for a *reconstruction system* $\{r_1, r_2, \dots\} \subset \mathcal{H}$ we define the reconstruction space $\mathcal{R} \subset \mathcal{H}$ as

$$\mathcal{R} = \overline{\text{span}}\{r_k : k \in \mathbb{N}\}$$

and, likewise, its finite dimensional version is defined as

$$\mathcal{R}_N = \text{span}\{r_1, \dots, r_N\}, \quad N \in \mathbb{N}.$$

We will assume the sampling system to form an orthonormal basis and the reconstruction system to be a frame, since this will exactly be our setup, see Section 4. A fully general presentation of generalized sampling can be found in [3, 6]. Further it is natural to assume a certain subspace condition, indeed, it is required that

$$\mathcal{R} \cap \mathcal{S}^\perp = \{0\} \quad \text{and} \quad \mathcal{R} + \mathcal{S} \text{ is closed.} \quad (2.2)$$

This guarantees a well posedness of the finite dimensional reconstruction problem, cf. Theorem 2.1. Indeed, the reconstruction problem is now formulated as follows: given a finite number of measurements $\langle f, s_1 \rangle, \dots, \langle f, s_M \rangle$ of some unknown $f \in \mathcal{H}$ we wish to determine a reconstruction $f_N \in \mathcal{R}_N$ such that $\|f - f_N\|$ is small and $f_N \rightarrow f$ as $N \rightarrow \infty$ fast. The following theorem leads to the generalized sampling reconstruction and proves its existence.

Theorem 2.1 ([6]). *Let \mathcal{S}_M and \mathcal{R}_N be as above and $P_{\mathcal{S}_M}$ be the following finite rank operator*

$$P_{\mathcal{S}_M} : \mathcal{H} \rightarrow \mathcal{S}_M$$

$$f \mapsto \sum_{k=1}^M \langle f, s_k \rangle s_k.$$

If (2.2) holds, then there exists an $M \in \mathbb{N}$ such that the system of equations

$$\langle P_{\mathcal{S}_M} f_{N,M}, r_j \rangle = \langle P_{\mathcal{S}_M} f, r_j \rangle, \quad j = 1, \dots, N \quad (2.3)$$

has a unique solution $f_{N,M} \in \mathcal{R}_N$. Moreover, the smallest $M \in \mathbb{N}$ such that the system is uniquely solvable is the least number $M \in \mathbb{N}$ so that

$$c_{N,M} := \inf_{\substack{f \in \mathcal{R}_N, \\ \|f\|=1}} \|P_{\mathcal{S}_M} f\| > 0. \quad (2.4)$$

Furthermore,

$$\|f - P_{\mathcal{R}_N} f\| \leq \|f - f_{N,M}\| \leq \frac{1}{c_{N,M}} \|f - P_{\mathcal{R}_N} f\|, \quad (2.5)$$

where $P_{\mathcal{R}_N} : \mathcal{H} \rightarrow \mathcal{R}_N$ denotes the orthogonal projection onto \mathcal{R}_N .

Definition 2.2 ([6]). *The solution $f_{N,M}$ in Theorem 2.1 is called generalized sampling reconstruction.*

Definition 2.3 ([34, 32]). *The quantity $c_{N,M}$ in (2.4) is called the infimum cosine angle between the subspaces \mathcal{R}_N and \mathcal{S}_M .*

2.2 Stable sampling rate

In light of Theorem 2.1 it is crucial to have control over $c_{N,M}$. This motivates the definition of the stable sampling rate.

Definition 2.4. *For any fixed $N \in \mathbb{N}$ and $\theta > 1$ the stable sampling rate $\Theta(N, \theta)$ is defined as*

$$\Theta(N, \theta) = \min \left\{ M \in \mathbb{N} : c_{N,M} > \frac{1}{\theta} \right\}.$$

The stable sampling rate determines the number of measurements M that are needed in order to find an approximation $f_{N,M} \in \mathcal{R}_N$ or equivalently, find N coefficients such that the angle $c_{N,M}$ is controlled by the threshold θ . We wish to emphasize that the control of $c_{N,M}$ allows the control of the error that is made by the approximation via GS by (2.5).

3 Shearlets

In this section we introduce the function systems that are used for the reconstruction space, these are *compactly supported shearlets*.

Shearlet systems were first introduced by K. Guo, G. Kutyniok, D. Labate, W.-Q. Lim and G. Weiss in [20, 26] and the following notations have become standard in this topic. The *parabolic scaling matrices* with respect to scale $j \in \mathbb{N} \cup \{0\}$ are denoted by

$$A_{2^j} := \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix}, \quad \tilde{A}_{2^j} := \begin{pmatrix} 2^{j/2} & 0 \\ 0 & 2^j \end{pmatrix},$$

and the *shearing matrices* with parameter $k \in \mathbb{Z}$ are

$$S_k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}.$$

These operations, together with the standard integer shift of functions in $L^2(\mathbb{R}^2)$ are then used to define the *cone adapted discrete shearlet system*.

Definition 3.1 ([22]). *Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be the generating functions and $c = (c_1, c_2) \in (\mathbb{R}^+)^2$. Then the (cone adapted discrete) shearlet system is defined as*

$$\mathcal{SH}(\phi, \psi, \tilde{\psi}, c) = \Phi(\phi, c_1) \cup \Psi(\psi, c) \cup \tilde{\Psi}(\tilde{\psi}, c),$$

where

$$\Phi(\phi, c_1) = \{\phi(\cdot - c_1 m) : m \in \mathbb{Z}^2\},$$

and

$$\begin{aligned} \Psi(\psi, c) &= \left\{ \psi_{j,k,m} = 2^{3j/4} \psi((S_k A_{2^j}) \cdot -cm) : j \geq 0, |k| \leq 2^{j/2}, m \in \mathbb{Z}^2 \right\}, \\ \tilde{\Psi}(\tilde{\psi}, c) &= \left\{ \tilde{\psi}_{j,k,m} = 2^{3j/4} \tilde{\psi}(S_k^T \tilde{A}_{2^j} \cdot -\tilde{c}m) : j \geq 0, |k| \leq 2^{j/2}, m \in \mathbb{Z}^2 \right\}, \end{aligned}$$

where the multiplication of c and \tilde{c} with the translation parameter m should be understood entry wise.

The shearlet system has a multiscale structure, however, it does not form a classical *multiresolution analysis* (MRA), see [12, 29] for a definition of an MRA. Moreover, it is still an open question whether there exists compactly supported orthonormal shearlet bases. Nevertheless, shearlets can form a frame.

Theorem 3.2 ([22]). *Let $\phi, \psi \in {}^2(\mathbb{R}^2)$ such that*

$$|\hat{\phi}(\xi_1, \xi_2)| \leq C_1 \min\{1, |\xi_1|^{-r}\} \min\{1, |\xi_2|^{-r}\}$$

and

$$|\hat{\psi}(\xi_1, \xi_2)| \leq C_2 \min\{1, |\xi_1|^\alpha\} \min\{1, |\xi_1|^{-r}\} \min\{1, |\xi_2|^{-r}\},$$

for some constants $C_1, C_2 > 0$ and $\alpha > r > 3$. Further let $\tilde{\psi}(x_1, x_2) = \psi(x_2, x_1)$ and assume there exists a positive constant $A > 0$ such that

$$|\hat{\phi}(\xi)|^2 + \sum_{j \geq 0} \sum_{|k| \leq \lceil 2^{j/2} \rceil} |\hat{\psi}(S_k^T (A_j)^{-1} \xi)|^2 + \sum_{j \geq 0} \sum_{|k| \leq \lceil 2^{j/2} \rceil} |\hat{\tilde{\psi}}(S_k (\tilde{A}_j)^{-1} \xi)|^2 > A$$

holds almost everywhere. Then there exists $c = (c_1, c_2) \in (\mathbb{R}^+)^2$ such that the cone-adapted shearlet system $\mathcal{SH}(\phi, \psi, \tilde{\psi}, c)$ forms a frame for $L^2(\mathbb{R}^2)$.

An explicit construction of compactly supported shearlets that form a frame can be found in [22]. Note that the frequency decay assumptions on the generators given in Theorem 3.2 are essential and secure that the overlapping in frequency domain is controlled. In particular, the frame constants depend on this behavior, [27, 22, 24]. For more information about shearlets we refer to [23] and the references therein.

3.1 Sparse approximation of cartoon-like functions

Shearlets are vastly used in applied harmonic analysis due to the *optimal sparse approximation rate for cartoon-like functions*. We now recall the definition of cartoon-like functions and the results about optimal sparse approximation rate.

The set of cartoon-like functions was first introduced in [13] and since then has been frequently used as a mathematical model for natural images.

Definition 3.3. Let $\nu > 0$ and $f : \mathbb{R}^2 \rightarrow \mathbb{C}$ a function of the form

$$f = g + h\chi_B,$$

where $B \subset [0, 1]^2$ is a set whose boundary is a closed C^2 curve with curvature bounded by ν and $g, h \in C^2(\mathbb{R})$ are compactly supported in $[0, 1]^2$ with $\|g\|_{C^2}, \|h\|_{C^2} \leq 1$. Then we call f a cartoon-like function. The set of cartoon-like functions is called class of cartoon-like functions and is denoted by $\mathcal{E}^2(\nu)$.

The following theorem shows the sparse approximation rate for shearlets.

Theorem 3.4 ([25]). Let $\mathcal{SH}(\phi, \psi, \tilde{\psi}, c)$ be a compactly supported shearlet frame with

$$i) |\hat{\psi}(\xi)| \leq C \min\{1, |\xi_1|^\alpha\} \min\{1, |\xi_1|^{-r}\} \min\{1, |\xi_2|^{-r}\},$$

$$ii) \left| \frac{\partial}{\partial \xi_2} \hat{\psi}(\xi) \right| \leq |h(\xi_1)| \left(1 + \frac{\xi_2}{\xi_1} \right)^{-r},$$

where $\alpha > 5, r \geq 4, h \in L^1(\mathbb{R}), C > 0$ is some constant. Further assume $\tilde{\psi}$ fulfills i) and ii) with switched roles of ξ_1 and ξ_2 . Then for any fixed $\nu > 0$ and $f \in \mathcal{E}^2(\nu)$ we have

$$\|f - f_N\|_2^2 \leq C' (\log N)^3 N^{-2}, \quad \text{as } N \rightarrow \infty,$$

where f_N is the best N -term approximation obtained using the N largest shearlets coefficients of f in magnitude allowing polynomial depth search and C' is some constant independent of N .

Remark 3.5. The rate obtained in Theorem 3.4 is also achieved in [?] for band-limited shearlet systems and for curvelets in [9].

In [13] it was shown, that under some assumptions on the representation system and the selection procedure of the largest coefficients in magnitude, the optimal approximation rate of a cartoon-like function f obeys

$$\|f - f_N\|^2 \leq CN^{-2}.$$

This rate is, up to a logarithmic factor, achieved by shearlets, cf. Theorem 3.4. Moreover, compared to the N^{-2} factor, the logarithmic term can be neglected and, thus, the rate in Theorem 3.4 is referred to as almost optimal sparse approximation rate.

Remark 3.6. It is commonly known, that standard wavelet orthonormal bases only obtain a rate of order N^{-1} , see [24]. Moreover both rates for shearlets and wavelets are sharp, that is, there exists cartoon-like functions such that the rate is achieved.

3.2 Assumptions on the generator

For the rest of this paper we assume ϕ and ψ to be compactly supported functions in $L^2(\mathbb{R}^2)$ and define $\tilde{\psi}(x_1, x_2) := \psi(x_2, x_1)$. Moreover, we assume the shearlets to have sufficient vanishing

moments and frequency decay, more precisely, we assume there exist some constants $C_1, C_2 > 0$ such that

$$|\widehat{\phi}(\xi_1, \xi_2)| \leq C_1 \cdot \frac{1}{(1 + |\xi_1|)^r} \frac{1}{(1 + |\xi_2|)^r}, \quad (3.1)$$

and

$$|\widehat{\psi}(\xi_1, \xi_2)| \leq C_2 \cdot \min\{1, |\xi_1|^\alpha\} \cdot \frac{1}{(1 + |\xi_1|)^r} \frac{1}{(1 + |\xi_2|)^r}. \quad (3.2)$$

where the regularity parameters $\alpha, r > 0$ are large enough so that the shearlet system forms a frame for $L^2(\mathbb{R}^2)$. These assumptions are equivalent to the frequency decay assumptions stated in Theorem 3.2.

4 Stable shearlet reconstructions from Fourier measurements

We will now define the reconstruction space and the sampling space for which we determine a stable sampling rate.

4.1 Shearlet reconstruction space

Without loss of generality we can assume that the generating scaling function ϕ and the shearlets $\psi, \tilde{\psi}$ are compactly supported in $[0, a]^2$ where a is some positive integer. Then we consider all scaling functions whose support intersect $[0, a]^2$ and denote this index set by Ω , i.e.

$$\Omega = \{m \in \mathbb{Z}^2 : \text{supp } \phi_m \cap [0, a]^2 \neq \emptyset\} = \{(m_1, m_2) \in \mathbb{Z}^2 : -a \leq m_1, m_2 \leq a\}.$$

Similarly, we consider all shearlets whose support intersect $[0, a]^2$. For this, let $J - 1 \in \mathbb{N} \cup \{0\}$ be a fixed scale. Then Λ_J denotes the following parameter set

$$\Lambda_J = \left\{ (j, k, m) \in \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}^2 : 0 \leq j \leq J - 1, |k| \leq 2^{j/2}, m \in \Omega_{j,k} \right\},$$

where $\Omega_{j,k} = \{m \in \mathbb{Z}^2 : \text{supp } \psi_{j,k,m} \cap [0, a]^2 \neq \emptyset\}$ is, due to the compact support of each $(\psi_{j,k,m})_m$, of finite cardinality and for $\tilde{\psi}$ and the second cone we analogously write

$$\tilde{\Lambda}_J = \left\{ (\tilde{j}, \tilde{k}, \tilde{m}) \in \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}^2 : 0 \leq \tilde{j} \leq J - 1, |\tilde{k}| \leq 2^{\tilde{j}/2}, \tilde{m} \in \tilde{\Omega}_{\tilde{j}, \tilde{k}} \right\}$$

with $\tilde{\Omega}_{\tilde{j}, \tilde{k}} = \{\tilde{m} \in \mathbb{Z}^2 : \text{supp } \tilde{\psi}_{\tilde{j}, \tilde{k}, \tilde{m}} \cap [0, a]^2 \neq \emptyset\}$ being of finite cardinality, respectively. In fact, the scale determines the number of translation and will be of order $2^{3/2j}$ for a fixed scale j . The reconstruction space at scale $J - 1$ is then defined as

$$\mathcal{R}_{N_J} := \text{span} \left\{ \{\phi_m : m \in \Omega\} \cup \{\psi_{j,k,m} : (j, k, m) \in \Lambda_J\} \cup \{\tilde{\psi}_{\tilde{j}, \tilde{k}, \tilde{m}} : (\tilde{j}, \tilde{k}, \tilde{m}) \in \tilde{\Lambda}_J\} \right\}. \quad (4.1)$$

Up to a fixed scale $J - 1$ we, asymptotically, have $N_J = 2^{2J}$ many generating functions in \mathcal{R}_{N_J} as $J \rightarrow \infty$. Note that by construction at each scale we only have finitely many elements, therefore, an ordering can be performed quite naturally namely we order the system along scales and within the scales, the shearing is ordered from $-2^{j/2}$ to $2^{j/2}$ and the translations in \mathbb{Z}^2 are ordered in a lexicographical manner for each fixed shearing and scaling parameter.

By \mathcal{R} we denote the reconstruction space that contains all shearlets across all scales, i.e.

$$\begin{aligned} \mathcal{R} = \overline{\text{span}} \left\{ \{\phi_m : m \in \Omega\} \cup \{\psi_{j,k,m} : (j, k, m) \in \Lambda_J, J \in \mathbb{N} \cup \{0\}\} \right. \\ \left. \cup \{\tilde{\psi}_{\tilde{j}, \tilde{k}, \tilde{m}} : (\tilde{j}, \tilde{k}, \tilde{m}) \in \tilde{\Lambda}_J, J \in \mathbb{N} \cup \{0\}\} \right\}. \end{aligned}$$

4.2 Fourier sampling space

To define the sampling space we first choose $T_1, T_2 > 0$ sufficiently large such that

$$\mathcal{R} \subset L^2([-T_1, T_2]^2).$$

Let $\varepsilon \leq \frac{1}{T_1+T_2} < 1$ control the sampling density. Then we define the sampling vectors on the uniform grid \mathbb{Z}^2 by

$$s_\ell^{(\varepsilon)} = \varepsilon e^{2\pi i \varepsilon \langle \ell, \cdot \rangle} \cdot \chi_{[-T_1, T_2]^2}, \quad \ell \in \mathbb{Z}^2. \quad (4.2)$$

Based on these sampling vectors we define the sampling space \mathcal{S} by

$$\mathcal{S}^{(\varepsilon)} = \overline{\text{span}} \left\{ s_\ell^{(\varepsilon)} : \ell \in \mathbb{Z}^2 \right\}.$$

For $M = (M_1, M_2) \in \mathbb{N} \times \mathbb{N}$ let

$$\mathcal{S}_M^{(\varepsilon)} = \text{span} \left\{ s_\ell^{(\varepsilon)} : \ell = (\ell_1, \ell_2) \in \mathbb{Z}^2, -M_i \leq \ell_i \leq M_i, i = 1, 2 \right\}$$

be the finite dimensional sampling space. Note that $M = (M_1, M_2)$ determines the size of the measured grid and the total number of possible measurements are in this case asymptotically of order $M_1 \cdot M_2$.

The task for us is now to describe the relationship between M and N in terms of the stable sampling rate which means such that stable and convergent reconstructions exist. In the event that the stable sampling rate is linear, i.e. $\Theta(N, \theta) = \mathcal{O}(N)$ we would have $M_1 \cdot M_2 = \mathcal{O}(2^{2J})$ as for the wavelet case, [7, 5].

4.3 Main results

Our main result shows that the angle between the shearlet reconstruction space and the Fourier sampling space can be controlled with an almost linear stable sampling rate.

Theorem 4.1. *Let $\mathcal{SH}(\phi, \psi, \tilde{\psi})$ be a compactly supported shearlet frame with generators ϕ, ψ , and $\tilde{\psi}$ and let $N \leq N_J = \mathcal{O}(2^{2J})$. Then for all $\theta > 1$ there exists $S_\theta > 0$ such that*

$$c_{N,M} = \inf_{\substack{f \in \mathcal{R}_N \\ \|f\|=1}} \|P_{\mathcal{S}_M^{(\varepsilon)}} f\| \geq \frac{1}{\theta},$$

where $M = (M_1, M_2) \in \mathbb{N} \times \mathbb{N}$ with $M_i = \lceil S_\theta A_N^{-1/(2r-1)} \cdot 2^{J(1+\delta)} / \varepsilon \rceil$, $\delta \geq \frac{2}{2r-1}$ and $r > 0$ is the regularity parameter from (3.1) and (3.2). Therefore $\Theta(N, \theta) = \mathcal{O}(N^{1+\delta} A_N^{-1/(2r-1)})$. Further, the constant S_θ does not depend on N but on θ, α , and r .

Remark 4.2. *The rate appearing in Theorem 4.1 is slightly worse than the rate for wavelets [7, 5] where the stable sampling rate is indeed linear, i.e. $\Theta(N, \theta) = \mathcal{O}(N)$. More precisely, one has to oversample by a factor N^δ and the rate depends on the lower frame bound A_N of the finite shearlet system that is used to span \mathcal{R}_N .*

Remark 4.3. *If we interpret the reciprocal of the lower frame bound as a measure for redundancy or stability, then our result states that more samples are necessary if the system is more redundant or less stable. Note that if the frame is tight, then the additional factor A_N^{-1} can be neglected.*

If one uses the same number of shearlets and wavelets to reconstruct a function from its Fourier measurements, then by Theorem 4.1, one has to acquire more Fourier measurements to guarantee the existence of the GS solution if shearlets are used compared to wavelets. This immediately raises the question why the use of shearlets in the context of GS is justified, if one has to acquire more samples to guarantee stable and convergent solutions. We address this issue from a different perspective. Suppose we were to sample an image using the rate computed in Theorem 4.1 for shearlets, but we want to reconstruct the image from the same number of measurements using shearlets and wavelets. Then, due to the stable sampling rate for shearlets and wavelets, we are allowed to use more wavelets than shearlets. More precisely, let $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ be the following oversampling function

$$\sigma(N) = \left\lceil N^{1+\delta} \frac{1}{A_N^{\frac{2}{2r-1}}} \right\rceil, \quad N \in \mathbb{N}. \quad (4.3)$$

Then the following result holds.

Corollary 4.4. *Let $\mathcal{SH}(\phi, \psi, \tilde{\psi})$ be a compactly supported shearlet frame with generators ϕ, ψ , and $\tilde{\psi}$ with sufficiently large regularity r and $\delta \geq 2/(2r-1)$. Further, let f be a cartoon-like function. For $N \in \mathbb{N}$ denote by f_N^s the best N -term approximation of f using shearlets and $f_{\sigma(N)}^w$ the best $\sigma(N)$ -approximation using wavelets, cf. Subsection 3.1. Further, let $N_J = 2^{2J}$, $J \in \mathbb{N}$ be the smallest number such that $N, \sigma(N) \lesssim N_J$. If $\mathcal{R}_{N_J}^s$ denotes the shearlet reconstruction space so that $f_N^s \in \mathcal{R}_{N_J}^s$ and $\mathcal{R}_{\sigma(N_J)}^w$ denotes the wavelet reconstruction space so that $f_{\sigma(N)}^w \in \mathcal{R}_{\sigma(N_J)}^w$, then*

$$\|f - G_{N_J, M}^s(f)\| \lesssim N^{-1}(\log N)^{3/2},$$

and

$$\|f - G_{\sigma(N_J), M'}^w(f)\| \lesssim \sigma(N)^{-1/2},$$

where M and M' are of order $\sigma(N_J)$ and $G_{N_J, M}^s(f)$ and $G_{\sigma(N_J), M'}^w(f)$ are the GS solution with respect to the spaces $\mathcal{R}_{N_J}^s$ and $\mathcal{R}_{\sigma(N_J)}^w$.

Remark 4.5. *Note that $N^{-1}(\log N)^{3/2} \lesssim \sigma(N)^{-1/2}$ if*

$$N^{-(1-\delta)/2}(\log N)^{3/2} \lesssim A_N^{1/(2r-1)}, \quad (4.4)$$

and therefore, if $\|f - f_N^s\| \asymp \|f - G_{N_J, M}^s(f)\|$ and $\|f - f_N^w\| \asymp \|f - G_{\sigma(N_J), M'}^w(f)\|$, then

$$\|f - G_{N_J, M}^s(f)\| \lesssim \|f - G_{\sigma(N_J), M'}^w(f)\|,$$

where " \asymp " means equality up to some constant.

In general, there is no control about the lower frame bound of the finite shearlet system known, thus, the reader might wonder whether (4.4) can hold. In Figure 2 in Section 6 we numerically confirm the validity of such an behavior of the lower frame bound.

5 Proofs

We first sketch the idea of the proof of Theorem 4.1.

5.1 Intuitive argument for the proof of Theorem 4.1

In order to prove the main theorem we have to bound

$$\inf_{\substack{f \in \mathcal{R}_N \\ \|f\|=1}} \|P_{\mathcal{S}_M^{(\varepsilon)}} f\|$$

from below for respective M and N . For any $f \in \mathcal{R}_N$ we have

$$\|P_{\mathcal{S}_M^{(\varepsilon)}} f\|^2 = \sum_{\ell \in I_M} |\langle f, s_\ell^{(\varepsilon)} \rangle|^2,$$

where $I_M = \{\ell = (\ell_1, \ell_2) \in \mathbb{Z}^2 : -M_i \leq \ell_i \leq M_i, i = 1, 2\}$. Hence, we could equivalently bound

$$\|P_{\mathcal{S}_M^{(\varepsilon)}}^\perp f\|^2 = \sum_{\ell \in (I_M)^c} |\langle f, s_\ell^{(\varepsilon)} \rangle|^2, \quad (5.1)$$

from above. Now, the key idea is to use the effective frequency support of the shearlets, in particular, the energy of $(\psi_{j,k,m})^\wedge$ is essentially localized in frequency bands of width 2^{2j} up to some constant, see Subsection 5.2 for precise statements. Moreover, since $f \in \mathcal{R}_N$ is a linear combination of shearlets up to a fixed scale $J-1$ the function f might be essentially supported in $[-C2^J, C2^J]^2$ for some constant C as well. Hence, by making the grid I_M large and dense enough the term in (5.1) should be small. However, there are two main concerns: First, the effective support could grow faster than linearly with respect to scaling factor 2^j . Second, taking linear combinations of shearlets might destroy the control of the frequency behavior and hence the essential supports in Fourier domain. Unfortunately, both obstacles are inevitable causing the additional factor N^δ and the dependence on the lower frame bound A_N in the stable sampling rate, respectively. More precisely, if $f \in \mathcal{R}_N$ then there exists $c_1, \dots, c_N \in \mathbb{C}$ such that

$$f = \sum_{\lambda \leq N} c_\lambda \psi_\lambda$$

but in general we have no control over the coefficients $(c_\lambda)_{\lambda \leq N}$. Therefore in order to resolve the latter concern, we expand f as follows:

$$f = \sum_{\lambda \leq N} \langle f, S_N^{-1} \psi_\lambda \rangle \psi_\lambda.$$

Then

$$\sum_{\lambda \leq N} |\langle f, S_N^{-1} \psi_\lambda \rangle|^2 \leq \frac{1}{A_N} \|f\|^2.$$

This is how the lower frame bound of the finite shearlet system comes into the proof and thus effects the stable sampling rate. Note that this problem does not appear if the reconstruction system is an orthonormal basis or a Riesz basis, thus, an additional penalty of the stable sampling rate is not necessary for wavelets.

5.2 Effective frequency support

The decay assumptions (3.1) and (3.2) give rise to the essential support of each shearlet atom ψ_λ in frequency, see Figure 1. The next result secures that the effective frequency support is controllable.

Proposition 5.1. *Let $J \in \mathbb{N}$ and $(\psi_\lambda)_{\lambda \leq N_J}$ be all shearlets up to scale $J - 1$ and $\omega > 0$. Then there exists a constant $S := S(\omega, r, \varepsilon)$ such that for $I_M := \{(\ell_1, \ell_2) \in \mathbb{Z}^2 : -M_i \leq \ell_i \leq M_i, i = 1, 2\}$ with $M_i = S2^{J(1+\delta)}, i = 1, 2$ and $\delta \geq \frac{2}{2r-1}$ we have*

$$\sum_{\ell \in (I_M)^c} \sum_{\lambda \leq N_J} |(\psi_\lambda)^\wedge(\varepsilon \ell)|^2 \leq \omega.$$



Figure 1: Effective frequency support of shearlets and tiling of the frequency plane, light regions correspond to large values, therefore regions with large energy and dark regions to small values, i.e. regions with small energy. Images are computed using the ShearLab package downloaded from <http://www.shearlab.org/>

To prove Proposition 5.1 we will use the following lemma which can be found in [17] (Appendix K).

Lemma 5.2 ([17]). *For $y \in \mathbb{R}, r > 1$, and $a, b > 0$ we have*

$$\int_0^\infty \frac{1}{(1 + a|x|)^r} \frac{1}{(1 + b|x - y|)^r} dx \lesssim \frac{1}{\max(a, b)} \frac{1}{(1 + \min(a, b)|y|)^r}.$$

Proof of Proposition 5.1. Let $\omega > 0$. Then we have to show the existence of an $M = (M_1, M_2) \in \mathbb{N} \times \mathbb{N}$, so that

$$\left(\sum_{|l_2| > M_2} \sum_{|l_1| > M_1} + \sum_{|l_2| < M_2} \sum_{|l_1| > M_1} + \sum_{|l_2| > M_2} \sum_{|l_1| < M_1} \right) \sum_{\lambda \leq N_J} |(\psi_\lambda)^\wedge(\varepsilon \ell)|^2 \leq \omega, \quad (5.2)$$

where M_1, M_2 scales like $S2^{J(1+\delta)}$ with a constant S independent of J sufficiently large. Note that

$$\begin{aligned} & \sum_{\lambda \leq N_J} |(\psi_\lambda)^\wedge(\varepsilon \ell)|^2 \\ &= \sum_{m' \in \Omega} \left| \varepsilon e^{-2\pi i \varepsilon \langle l, m' \rangle} \widehat{\phi}(\varepsilon l) \right|^2 + \sum_{(j, k, m) \in \Lambda_J} \left| \frac{\varepsilon}{2^{3j/4}} e^{-2\pi i \varepsilon \langle (S_k A_{2j})^{-T} l, m \rangle} \widehat{\psi} \left(\varepsilon (S_k A_{2j})^{-T} l \right) \right|^2 + \\ & \quad + \sum_{(\tilde{j}, \tilde{k}, \tilde{m}) \in \tilde{\Lambda}_J} \left| \frac{\varepsilon}{2^{3\tilde{j}/4}} e^{-2\pi i \varepsilon \langle (S_{\tilde{k}} \tilde{A}_{2\tilde{j}})^{-T} l, \tilde{m} \rangle} \widehat{\tilde{\psi}} \left(\varepsilon (S_{\tilde{k}} \tilde{A}_{2\tilde{j}})^{-T} l \right) \right|^2, \end{aligned}$$

$$\leq C \left| \varepsilon \widehat{\phi}(\varepsilon l) \right|^2 + \sum_{j=0}^{J-1} \sum_{k=-2^{j/2}}^{2^{j/2}} \left| \widehat{\psi} \left(\varepsilon (S_k A_{2^j})^{-T} l \right) \right|^2 + \sum_{\tilde{j}=0}^{J-1} \sum_{\tilde{k}=-2^{\tilde{j}/2}}^{2^{\tilde{j}/2}} \left| \widehat{\psi} \left(\varepsilon (S_{\tilde{k}} \tilde{A}_{2^{\tilde{j}}})^{-T} l \right) \right|^2, \quad (5.3)$$

since the cardinality of $\Omega_{j,k}$ and $\tilde{\Omega}_{\tilde{j},\tilde{k}}$ is of order $2^{3/2j}$.

Denote by I, II, III the following terms

$$\begin{aligned} \text{I} &:= \left(\sum_{|l_2| > M_2} \sum_{|l_1| > M_1} + \sum_{|l_2| < M_2} \sum_{|l_1| > M_1} + \sum_{|l_2| > M_2} \sum_{|l_1| < M_1} \right) \left| \varepsilon \widehat{\phi}(\varepsilon l) \right|^2, \\ \text{II} &:= \left(\sum_{|l_2| > M_2} \sum_{|l_1| > M_1} + \sum_{|l_2| < M_2} \sum_{|l_1| > M_1} + \sum_{|l_2| > M_2} \sum_{|l_1| < M_1} \right) \sum_{j=0}^{J-1} \sum_{k=-2^{j/2}}^{2^{j/2}} \left| \widehat{\psi} \left(\varepsilon (S_k A_{2^j})^{-T} l \right) \right|^2, \\ \text{III} &:= \left(\sum_{|l_2| > M_2} \sum_{|l_1| > M_1} + \sum_{|l_2| < M_2} \sum_{|l_1| > M_1} + \sum_{|l_2| > M_2} \sum_{|l_1| < M_1} \right) \sum_{\tilde{j}=0}^{J-1} \sum_{\tilde{k}=-2^{\tilde{j}/2}}^{2^{\tilde{j}/2}} \left| \widehat{\psi} \left(\varepsilon (S_{\tilde{k}} \tilde{A}_{2^{\tilde{j}}})^{-T} l \right) \right|^2. \end{aligned}$$

We will estimate each sum by using the decay conditions (3.1) and (3.2), respectively, in order to obtain (5.2) for sufficiently large S independent on J with $M_i = \frac{2^{J(1+\delta)}}{\varepsilon} S, i = 1, 2$. Since

$$\left| \varepsilon \widehat{\phi}(\varepsilon l) \right|^2 \leq C_1^2 \varepsilon^2 \left| \frac{1}{(1 + |\varepsilon l_1|)^r} \frac{1}{(1 + |\varepsilon l_2|)^r} \right|^2. \quad (5.4)$$

we immediately have

$$\begin{aligned} \sum_{|l_2| > M_2} \sum_{|l_1| < M_1} \left| \widehat{\phi}(\varepsilon l) \right|^2 &\leq \sum_{|l_2| > M_2} \sum_{|l_1| < M_1} C_1^2 \frac{1}{(1 + |\varepsilon l_1|)^{2r}} \frac{1}{(1 + |\varepsilon l_2|)^{2r}} \\ &\leq C_1^2 \sum_{|l_1| > M_1} \frac{1}{(1 + |\varepsilon l_1|)^{2r}} \\ &\leq C_1^2 \frac{1}{(1 + S^{J(1+\delta)})^{2r-1}}, \end{aligned} \quad (5.5)$$

where the constant C_1 changed in each step. Analogously, one obtains

$$\begin{aligned} \sum_{|l_2| < M_2} \sum_{|l_1| > M_1} \left| \widehat{\phi}(\varepsilon l) \right|^2 &\leq C_1^2 \frac{1}{(1 + S^{J(1+\delta)})^{2r-1}}, \\ \sum_{|l_2| > M_2} \sum_{|l_1| > M_1} \left| \widehat{\phi}(\varepsilon l) \right|^2 &\leq C_1^2 \frac{1}{(1 + S^{J(1+\delta)})^{2r-1}}. \end{aligned} \quad (5.6)$$

Hence, combining (5.5) and (5.6) gives

$$\text{I} \leq C_1^2 \frac{1}{(1 + S^{J'(1+\delta)})^{2r-1}} \quad (5.7)$$

Regarding II we first have by (3.2)

$$\sum_{j=0}^{J-1} \sum_{k=-2^{j/2}}^{2^{j/2}} \left| \widehat{\psi} \left(\varepsilon (S_k A_{2^j})^{-T} l \right) \right|^2 \leq \sum_{j=0}^{J-1} \sum_{k=-2^{j/2}}^{2^{j/2}} \frac{C_2^2}{(1 + |\varepsilon 2^{-j} l_1|)^{2r} (1 + |-\varepsilon k 2^{-j} l_1 + \varepsilon 2^{-j/2} l_2|)^{2r}}, \quad (5.8)$$

and, likewise for III,

$$\sum_{\tilde{j}=0}^{J-1} \sum_{\tilde{k}=-2^{\tilde{j}/2}}^{2^{\tilde{j}/2}} \left| \widehat{\psi} \left(\varepsilon \left(S_k^T A_{2^{\tilde{j}}} \right)^{-T} l \right) \right|^2 \leq \sum_{\tilde{j}=0}^{J-1} \sum_{\tilde{k}=-2^{\tilde{j}/2}}^{2^{\tilde{j}/2}} \frac{C_3^2}{(1 + |\varepsilon 2^{-\tilde{j}} l_2|)^{2r} (1 + |\varepsilon 2^{-\tilde{j}/2} l_1 - \varepsilon \tilde{k} 2^{-\tilde{j}} l_2|)^{2r}}. \quad (5.9)$$

We only continue to estimate (5.8) since an estimate for (5.9) is then obtained analogously.

We consider two cases. The first case concerns shearlets, that are *wavelet-like*. More precisely, we consider shearlets with no shearing first, so parabolically scaled wavelets.

Case I: Let $0 \leq j \leq J-1, k = 0$. By direct computations similar to (5.5) we have

$$\begin{aligned} & \sum_{|l_2| > M_2} \sum_{|l_1| < M_1} \sum_{j=0}^{J-1} C_2^2 \frac{1}{(1 + |\varepsilon 2^{-j} l_1|)^{2r}} \frac{1}{(1 + |\varepsilon 2^{-j/2} l_2|)^{2r}} \\ & \leq \sum_{j=0}^{J-1} C_2^2 \sum_{|l_2| > M_2} \frac{1}{(1 + |\varepsilon 2^{-j/2} l_2|)^{2r}} \sum_{|l_1| < M_1} \frac{1}{(1 + |\varepsilon 2^{-j} l_1|)^{2r}} \\ & \leq \sum_{j=0}^{J-1} C_2^2 \frac{2^j}{\varepsilon} \sum_{|l_2| > M_2} \frac{1}{(1 + |\varepsilon 2^{-j/2} l_2|)^{2r}} \\ & \leq \sum_{j=0}^{J-1} C_2^2 \frac{2^{3/2j}}{\varepsilon^2} \frac{1}{(1 + S 2^{J'(1+\delta)-j/2})^{2r-1}} \\ & \leq C_2^2 \frac{2^{3/2J}}{\varepsilon^2} \frac{1}{(1 + S 2^{J(1/2+\delta)})^{2r-1}}. \end{aligned} \quad (5.10)$$

where C_2 changed over time. In the same manner we obtain

$$\sum_{|l_2| < M_2} \sum_{|l_1| > M_1} \sum_{j=0}^{J-1} C_2^2 \frac{1}{(1 + |\varepsilon 2^{-j} l_1|)^{2r}} \frac{1}{(1 + |\varepsilon 2^{-j/2} l_2|)^{2r}} \leq C_2^2 \frac{2^{3/2J}}{\varepsilon^2} \frac{1}{(1 + S 2^{J\delta})^{2r-1}}, \quad (5.11)$$

and

$$\begin{aligned} & \sum_{|l_1| > M_1} \sum_{|l_2| > M_2} \sum_{j=0}^{J-1} C_2^2 \frac{1}{(1 + |\varepsilon 2^{-j} l_1|)^{2r}} \frac{1}{(1 + |\varepsilon 2^{-j/2} l_2|)^{2r}} \\ & \leq C_2^2 \frac{2^{3/2J}}{\varepsilon^2} \frac{1}{((1 + S 2^{J\delta})(1 + S 2^{J(1/2+\delta)}))^{2r-1}}. \end{aligned} \quad (5.12)$$

Case 2: Let $0 \leq j \leq J-1, k \neq 0$. By using Lemma 5.2 we have

$$\begin{aligned} & \sum_{|l_2| > M_2} \sum_{|l_1| < M_1} \sum_{j=0}^{J-1} \sum_{\substack{k=-2^{j/2} \\ k \neq 0}}^{2^{j/2}} C_2^2 \frac{1}{(1 + |\varepsilon 2^{-j} l_1|)^{2r}} \frac{1}{(1 + |-\varepsilon k 2^{-j} l_1 + \varepsilon 2^{-j/2} l_2|)^{2r}} \\ & \leq \sum_{j=0}^{J-1} \sum_{\substack{k=-2^{j/2} \\ k \neq 0}}^{2^{j/2}} \sum_{|l_2| > M_2} C_2^2 \frac{1}{\varepsilon |k 2^{-j}|} \frac{1}{(1 + |\varepsilon 2^{-j}| |2^{j/2} l_2 / k|)^{2r}} \\ & \leq \sum_{j=0}^{J-1} \sum_{\substack{k=-2^{j/2} \\ k \neq 0}}^{2^{j/2}} C_2^2 \frac{2^j}{\varepsilon |k|} \frac{|k| 2^{j/2}}{\varepsilon} \frac{1}{(1 + |\frac{S}{\varepsilon} 2^{J'(1+\delta)-j/2} / k|)^{2r-1}} \end{aligned}$$

$$\begin{aligned}
&= \sum_{j=0}^{J-1} \sum_{\substack{k=-2^{j/2} \\ k \neq 0}}^{2^{j/2}} C_2^2 \frac{2^{3/2j}}{\varepsilon^2} \frac{1}{(1 + |\frac{S}{\varepsilon} 2^{J(1+\delta)-j/2}/k|)^{2r-1}} \\
&\leq C_2^2 \frac{2^{2J}}{\varepsilon^2} \frac{1}{(1 + |\frac{S}{\varepsilon} 2^{J\delta}|)^{2r-1}}.
\end{aligned} \tag{5.13}$$

Since

$$\sum_{|l_2| < M_2} \frac{1}{(1 + |-\varepsilon k 2^{-j} l_1 + \varepsilon 2^{-j/2} l_2|)^{2r}} \lesssim \int_{\mathbb{R}} \frac{1}{(1 + |-\varepsilon k 2^{-j} l_1 + \varepsilon 2^{-j/2} x|)^{2r}} dx \lesssim \frac{2^{j/2}}{\varepsilon}$$

we have that

$$\begin{aligned}
&\sum_{|l_2| < M_2} \sum_{|l_1| > M_1} \sum_{j=0}^{J-1} \sum_{\substack{k=-2^{j/2} \\ k \neq 0}}^{2^{j/2}} C_2^2 \frac{1}{(1 + |\varepsilon 2^{-j} l_1|)^{2r}} \frac{1}{(1 + |-\varepsilon k 2^{-j} l_1 + \varepsilon 2^{-j/2} l_2|)^{2r}} \\
&\leq \sum_{j=0}^{J-1} \sum_{\substack{k=-2^{j/2} \\ k \neq 0}}^{2^{j/2}} C_2^2 \frac{2^{3/2j}}{\varepsilon^2} \frac{1}{(1 + |\frac{S}{\varepsilon} 2^{J(1+\delta)-j}|)^{2r-1}} \\
&\leq C_2^2 \frac{2^{2J}}{\varepsilon^2} \frac{1}{(1 + |\frac{S}{\varepsilon} 2^{J\delta}|)^{2r-1}}.
\end{aligned} \tag{5.14}$$

Finally, the last sum can be bounded as in (5.13) by first using Lemma 5.2 again

$$\begin{aligned}
&\sum_{|l_2| > M_2} \sum_{|l_1| > M_1} \sum_{j=0}^{J-1} \sum_{\substack{k=-2^{j/2} \\ k \neq 0}}^{2^{j/2}} C_2^2 \frac{1}{(1 + |\varepsilon 2^{-j} l_1|)^{2r}} \frac{1}{(1 + |-\varepsilon k 2^{-j} l_1 + \varepsilon 2^{-j/2} l_2|)^{2r}} \\
&\leq \sum_{j=0}^{J-1} \sum_{\substack{k=-2^{j/2} \\ k \neq 0}}^{2^{j/2}} C_2^2 \frac{2^{3/2j}}{\varepsilon^2} \frac{1}{(1 + |S_\theta 2^{J(1+\delta)-j}|)^{2r-1}} \\
&\leq C_2^2 \frac{2^{2J}}{\varepsilon^2} \frac{1}{(1 + |\frac{S}{\varepsilon} 2^{J\delta}|)^{2r-1}}.
\end{aligned} \tag{5.15}$$

hence, combining (5.10), (5.11), (5.12), (5.13) (5.14), (5.15) yields

$$\text{II} \leq C_2^2 \frac{2^{2J}}{\varepsilon^2} \frac{1}{(1 + |\frac{S}{\varepsilon} 2^{J\delta}|)^{2r-1}}. \tag{5.16}$$

Regarding III, we obtain a similar estimate by performing the same computations as for II, therefore

$$\text{III} \leq C_2^2 \frac{2^{2J}}{\varepsilon^2} \frac{1}{(1 + |\frac{S}{\varepsilon} 2^{J\delta}|)^{2r-1}}. \tag{5.17}$$

Combining (5.7), (5.16), (5.17) yields

$$\text{I} + \text{II} + \text{III} \leq C \frac{\varepsilon^{2r-3}}{(S 2^{J(\delta-2/(2r-1))})^{2r-1}},$$

with a constant that does not depend on J . Therefore, if $\delta \geq \frac{2}{2r-1}$ the result follows for S greater than $C^{-1}(\frac{\varepsilon^{2r-3}}{\omega})^{1/(2r-1)}$. \square

5.3 Proof of Theorem 4.1

The proof of Theorem 4.1 is a consequence of Proposition 5.1.

Proof of Theorem 4.1. Let $\theta > 1$. Then we want to show

$$\inf_{\substack{f \in \mathcal{R}_N \\ \|f\|=1}} \|P_{S_M^{(\varepsilon)}} f\| \geq \frac{1}{\theta} \quad (5.18)$$

for appropriate M . For this, let $f \in \mathcal{R}_N$ with $\|f\| = 1$. We prove

$$\|P_{S_M^{(\varepsilon)}}^\perp f\|^2 \leq \frac{\theta^2 - 1}{\theta^2}, \quad (5.19)$$

for the claimed M . Since $(s_\ell)_\ell$ is an orthonormal system, we have

$$\|P_{S_M^{(\varepsilon)}}^\perp f\|^2 = \sum_{l \in (I_M)^c} |\langle f, s_l^{(\varepsilon)} \rangle|^2,$$

where $(I_M)^c$ denotes the set complement of I_M in \mathbb{Z}^2 , in particular

$$\sum_{l \in (I_M)^c} |\langle f, s_l^{(\varepsilon)} \rangle|^2 = \sum_{|l_2| > M_2} \sum_{|l_1| > M_1} |\langle f, s_l^{(\varepsilon)} \rangle|^2 + \sum_{|l_2| < M_2} \sum_{|l_1| > M_1} |\langle f, s_l^{(\varepsilon)} \rangle|^2 + \sum_{|l_2| > M_2} \sum_{|l_1| < M_1} |\langle f, s_l^{(\varepsilon)} \rangle|^2.$$

Therefore we have

$$f = \sum_{\lambda \leq N_J} \langle f, S_N^{-1} \psi_\lambda \rangle \psi_\lambda = \sum_{m' \in \Omega} \alpha_{m'} \phi_{m'} + \sum_{(j,k,m) \in \Lambda_J} \beta_{j,k,m} \psi_{j,k,m} + \sum_{(\tilde{j}, \tilde{k}, \tilde{m}) \in \tilde{\Lambda}_J} \gamma_{\tilde{j}, \tilde{k}, \tilde{m}} \tilde{\psi}_{\tilde{j}, \tilde{k}, \tilde{m}}.$$

with

$$\alpha_{m'} = \langle \phi_{m'}, s_l \rangle, \quad \beta_{j,k,m} = \langle \psi_{j,k,m}, s_l \rangle, \quad \gamma_{\tilde{j}, \tilde{k}, \tilde{m}} = \langle \tilde{\psi}_{\tilde{j}, \tilde{k}, \tilde{m}}, s_l \rangle.$$

Furthermore,

$$\sum_{m' \in \Omega} |\alpha_{m'}|^2 + \sum_{(j,k,m) \in \Lambda_J} |\beta_{j,k,m}|^2 + \sum_{(\tilde{j}, \tilde{k}, \tilde{m}) \in \tilde{\Lambda}_J} |\gamma_{\tilde{j}, \tilde{k}, \tilde{m}}|^2 = \sum_{\lambda \leq N_J} |\langle S_N^{-1} f, \psi_\lambda \rangle|^2 \leq \frac{1}{A_N} \|f\|^2.$$

Therefore by Cauchy Schwarz

$$\begin{aligned} & \sum_{l \in (I_M)^c} |\langle f, s_l \rangle|^2 \\ &= \sum_{l \in (I_M)^c} \left| \left\langle \sum_{m' \in \Omega} \alpha_{m'} \phi_{m'} + \sum_{(j,k,m) \in \Lambda_J} \beta_{j,k,m} \psi_{j,k,m} + \sum_{(\tilde{j}, \tilde{k}, \tilde{m}) \in \tilde{\Lambda}_J} \gamma_{\tilde{j}, \tilde{k}, \tilde{m}} \tilde{\psi}_{\tilde{j}, \tilde{k}, \tilde{m}}, s_l \right\rangle \right|^2 \\ &= \sum_{l \in (I_M)^c} \left| \sum_{m' \in \Omega} \alpha_{m'} \widehat{\phi_{m'}}(\varepsilon l) + \sum_{(j,k,m) \in \Lambda_J} \beta_{j,k,m} \widehat{\psi_{j,k,m}}(\varepsilon l) + \sum_{(\tilde{j}, \tilde{k}, \tilde{m}) \in \tilde{\Lambda}_J} \gamma_{\tilde{j}, \tilde{k}, \tilde{m}} \widehat{\tilde{\psi}_{\tilde{j}, \tilde{k}, \tilde{m}}}(\varepsilon l) \right|^2 \\ &\leq \sum_{l \in (I_M)^c} \frac{1}{A_N} \left(\sum_{m' \in \Omega} \left| \widehat{\phi_{m'}}(\varepsilon l) \right|^2 + \sum_{(j,k,m) \in \Lambda_J} \left| \widehat{\psi_{j,k,m}}(\varepsilon l) \right|^2 + \sum_{(\tilde{j}, \tilde{k}, \tilde{m}) \in \tilde{\Lambda}_J} \left| \widehat{\tilde{\psi}_{\tilde{j}, \tilde{k}, \tilde{m}}}(\varepsilon l) \right|^2 \right). \end{aligned}$$

By Proposition 5.1 there exists $S(\theta, r, \varepsilon)$ so that we can choose

$$M_i = \left\lceil S(\theta, r, \varepsilon) 2^{J(1+\delta)} A_N^{\frac{1}{2r-1}} \right\rceil \in \mathbb{N}, \quad i = 1, 2$$

to conclude

$$\sum_{l \in (I_M)^c} |\langle f, s_l \rangle|^2 \leq \frac{\theta^2 - 1}{\theta^2}$$

which shows (5.19) and hence

$$\|P_{S_M^\varepsilon} f\| \geq \frac{1}{\theta}.$$

□

5.4 Proof of Corollary 4.4

For the proof of Corollary 4.4 we recall the following results. By Theorem 3.4 the best N -term approximation f_N^s for shearlets of cartoon-like functions obeys

$$\|f - f_N^s\| \lesssim N^{-1}(\log N)^{3/2},$$

and for wavelets the best N -term approximation f_N^w obeys

$$\|f - f_N^w\| \lesssim N^{-1/2},$$

see [23]. Furthermore, these bounds are sharp, i.e. there exist cartoon-like functions such that the above inequalities are becoming equalities up to a constant.

Now, let $N \in \mathbb{N}$, $f_N^s, f_{\sigma(N)}^w, \mathcal{R}_{N_J}^s, \mathcal{R}_{\sigma(N_J)}^w$ be as in the corollary. By Theorem 4.1 there exist for fixed $\theta > 1$ an $M = (M_1, M_2)$ and $M' = (M'_1, M'_2)$ in $\mathbb{N} \times \mathbb{N}$ such that

$$c_{N_J, M} = \inf_{\substack{f \in \mathcal{R}_{N_J}^s \\ \|f\|=1}} \|P_{S_M^{(\varepsilon)}} f\| \geq \frac{1}{\theta}$$

and

$$c_{\sigma(N_J), M'} = \inf_{\substack{f \in \mathcal{R}_{\sigma(N_J)}^w \\ \|f\|=1}} \|P_{S_{M'}^{(\varepsilon)}} f\| \geq \frac{1}{\theta}.$$

Moreover, due to Theorem 4.1 and the linear stable sampling rate for wavelets proven in [7, 5] we can choose $M_1 \cdot M_2$ and $M'_1 \cdot M'_2$ to be of order $\sigma(N_J)$.

Using Theorem 2.1 we obtain

$$\|f - G_{\sigma(N_J), M'}^w(f)\| \lesssim \|f - P_{\mathcal{R}_{\sigma(N_J)}^w}(f)\| \lesssim \|f - f_{\sigma(N)}^w\| \lesssim \sigma(N)^{-1/2}.$$

Analogously, for shearlets we obtain using Theorem 2.1

$$\|f - G_{N_J, M}^s(f)\|^2 \lesssim \|f - P_{\mathcal{R}_{N_J}^s}(f)\|^2 \lesssim \left\| f - \sum_{\lambda \in I_N} \langle f, \psi_\lambda^d \rangle \psi_\lambda \right\|^2 \lesssim \sum_{\lambda \notin I_N} |\langle f, \psi_\lambda^d \rangle|^2,$$

where $(\psi_\lambda^d)_\lambda$ is the dual shearlet system of $(\psi_\lambda)_\lambda$ and I_N denotes the index set of the N largest coefficients $(\langle f, \psi_\lambda \rangle)_\lambda$ with respect to the best N -term approximation f_N^s .

Furthermore, if the regularity is sufficiently large, it was shown in [18] that we can relate the analysis coefficients of the primal frame with those of the dual frame by the relation

$$(\langle f, \psi_\lambda \rangle)_\lambda = G^\dagger(\langle f, \psi_\lambda^d \rangle)_\lambda,$$

where G^\dagger is the pseudo inverse of the Gramian operator associated to the shearlet system (ψ_λ) . Therefore the claimed rate follows from the decay rate of the shearlet coefficients. Hence,

$$\|f - G_{N_J, M}^s(f)\|^2 \lesssim \sum_{\lambda \notin I_N} |\langle f, \psi_\lambda \rangle|^2 \lesssim N^{-2}(\log N)^3$$

which yields the result.

6 Numerics

6.1 Inequality (4.4)

In light of Corollary 4.4 we want to compare the bounds $N^{-1}(\log N)^{3/2}$ and $\sigma(N)^{-1/2}$ with σ defined as in (4.3) for $N \in \mathbb{N}$. We therefore aim to verify (4.4) by showing the following behavior of the lower frame bound A_N

$$N^{-(1-\delta)/2}(\log N)^{3/2} \lesssim A_N^{1/(2r-1)}. \quad (6.1)$$

For convenience we will check (6.1) in levels, i.e. the lower frame bounds A_N will always belong to full shearlet systems at a scale J . Further, we use the SHEARLAB package which is available at

<http://www.shearlab.org>

to generate shearlet systems with respect to different levels and their respective lower frame bounds. We wish to mention that both objects are available in the implementation.

Now, note that a shearlet system consists of approximately 2^{2J} many elements for a fixed scale J where we ignore constants that are independent of J . We will numerically test (4.4) by plotting the lower frame bounds of a full shearlet system for increasing J . This yields a monotonically decreasing curve as the scale increase. This will now be compared to the decay of the function $f(N) = N^{-(1-\delta)/2}(\log N)^{3/2}$ with $N = 2^{2J}$ for different values of r and δ and for increasing J .

The informed reader might notice, that choosing $N = 2^{2J}$ is significantly moderate, since the constant which determines the number of shearlets at scale J is in practice much larger than 1. However, for larger N the function f decreases even faster. Moreover, since we are only interested in an asymptotic behavior of the lower frame bound with respect to increasing scales this is not a restriction. Indeed, already for that case it is strongly indicated that (4.4) hold asymptotically as depicted in Figure 2.

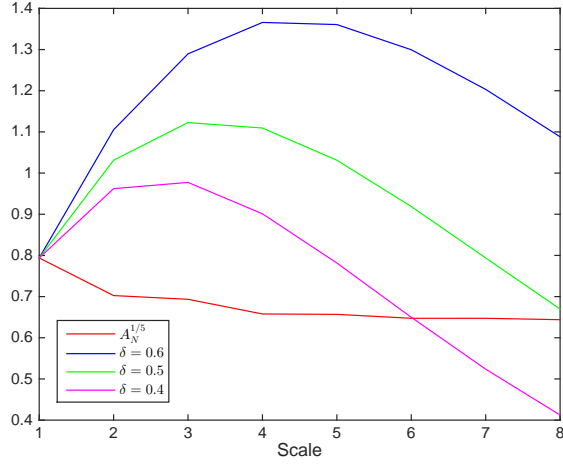
6.2 Reconstruction from MR data using shearlets

We further provide some reconstructions from MR data using complex exponentials (Fourier inversion), compactly supported Daubechies wavelets and compactly supported shearlets. Although our main result guarantees stable and convergent reconstructions when the sampling rate is almost linear it is not efficient (and also not necessary) to acquire that many samples in practice. In our numerics we will subsample the Fourier data to achieve practical relevance. All computation are done in MATLAB.

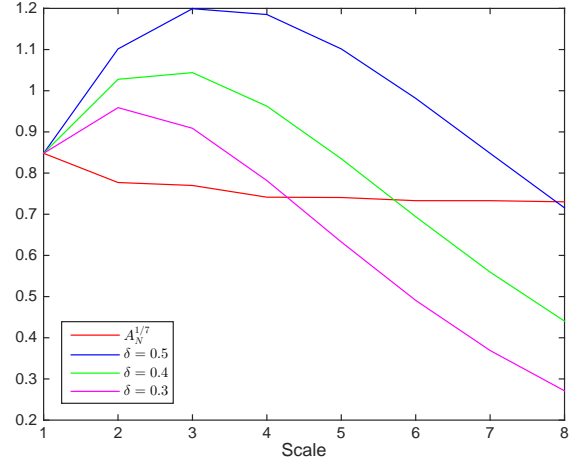
6.3 Data

The underlying Fourier data or also called *k-space* was acquired using a multi-channel acquisition consisting of four channels. Each of the four k-spaces has a 128×128 image resolution. Moreover, each channel data results in a single image of same size, e.g. by applying an inverse Fourier transform, see Figure 3a and Figure 3b.

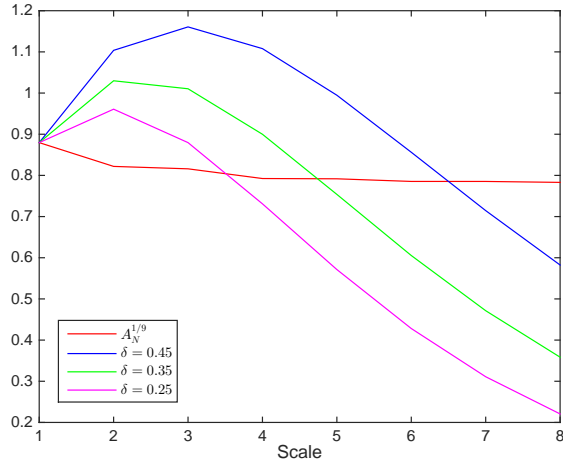
The single channel images are combined by using the sum-of-squares method [31]. The resulting sum-of-squares image from all four channels of the original k-space is used as the reference image, cf. Figure 3c.



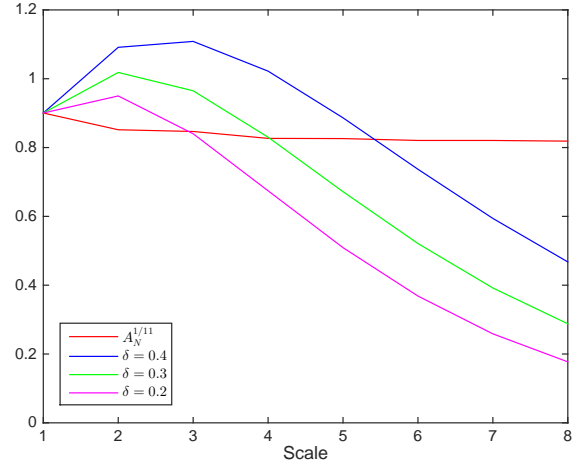
(a) $r = 3$ and $\delta = 0.4, 0.5, 0.6$



(b) $r = 4$ and $\delta = 0.3, 0.4, 0.5$

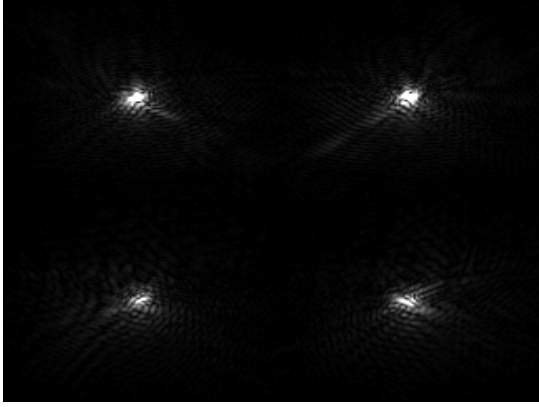


(c) $r = 5$ and $\delta = 0.25, 0.35, 0.45$

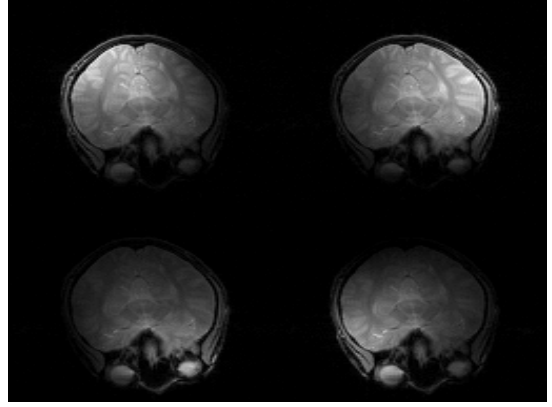


(d) $r = 6$ and $\delta = 0.2, 0.3, 0.4$

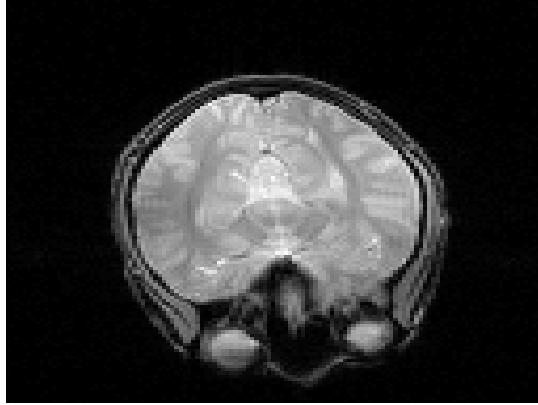
Figure 2: Plot of the behavior of lower frame bound and the critical function $N^{-(1-\delta)/2}(\log N)^{3/2}$ for different values of r and δ



(a) 4 channels input data



(b) 4 channel reconstructions



(c) Combined image by sum-of-squares

Figure 3: Reference images computed with complete data

6.4 Methods

We mimic the acceleration of factor 5 by using a 0–1 mask to subsample the k-space, cf. Figure 4a and Figure 4b. We have chosen a phyllotaxis spiral which is also used in [35] and a standard radial subsampling scheme.

The final reconstruction is obtained by computing each channel image separately and then combine it by using the *sum-of-squares method* [31]. More precisely, if I_k denotes the k -th image reconstructed from the k -th k-space then the final image I is computed by

$$I = \sqrt{|I_1| + \dots + |I_N|^2},$$

where N is the total number of channels and the the square operation, as well as the modulus is to be understood entry wise for the matrices I_k .

The reconstruction of the single channel images are computed as follows.

- Shearlets:

The reconstructions are obtained by solving the following unconstrained minimization problem

$$\min_u \frac{1}{2} \|\tilde{\mathcal{F}}u - b\|_2^2 + \lambda \|\mathcal{SH}(u)\|_1,$$

where $\tilde{\mathcal{F}}$ is the subsampled Fourier operator, b is the subsampled input data, \mathcal{SH} is the shearlet transform and λ are some parameters larger than zero, and u is the image.

The penalty function $\|\mathcal{SH}(u)\|_1$ enforces sparsity with respect to the shearlet system. The code for the shearlet transform is downloadable at

<http://www.shearlab.org>

- Wavelets:

Reconstructions are computed using the SparseMRI package from Lustig et al. [28] which is downloaded from

<http://www.eecs.berkeley.edu/~mlustig/Software.html>

This method is a standard tool in MRI to compute wavelet sparsity-based reconstructions from Fourier measurements.

- Fourier:

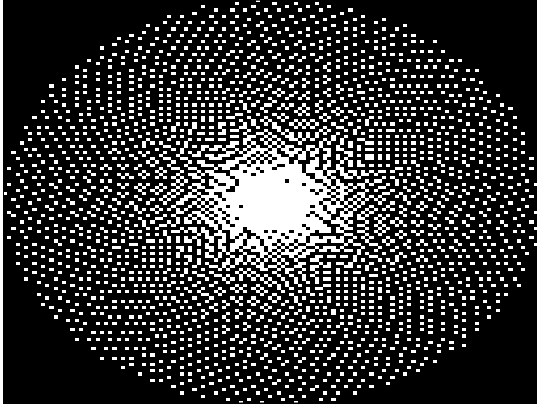
The reconstructions using complex exponentials are computed by taking the Fourier inversion of the subsampled data.

6.5 Results

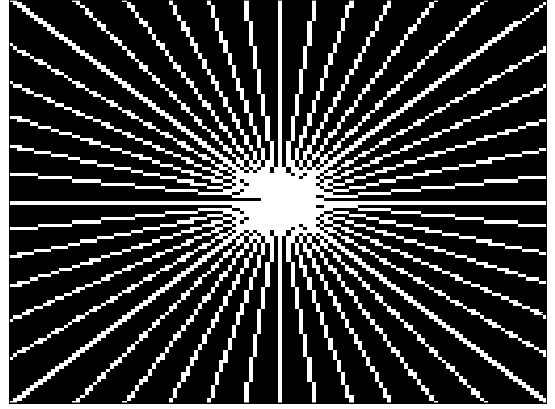
The reconstructions show less artifacts and improved image quality using shearlets compared to the reconstruction obtained from compactly supported wavelets, cf. Figure 4. Although the theory of optimal sparse approximation rates hold in a continuous setup, it is surprising to see such strong differences for images of such low resolutions such as 128×128 .

	Shearlets	Wavelets	Fourier inversion
Spiral mask	0.638	0.1006	0.2570
Radial mask	0.0847	0.1282	0.2690

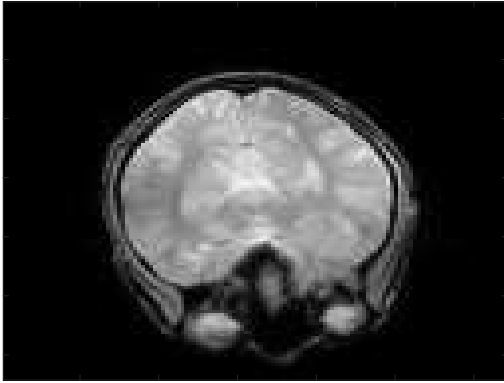
Table 1: Relative error for Figure 4



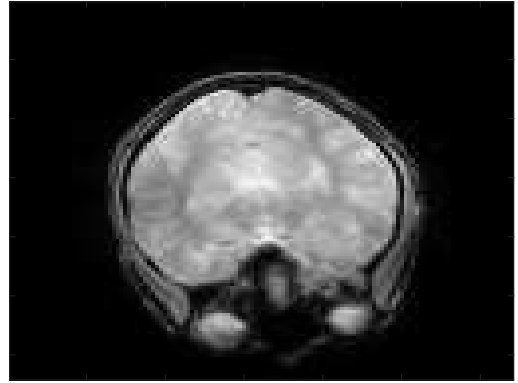
(a) Phyllotaxis spiral mask: 20.37 %



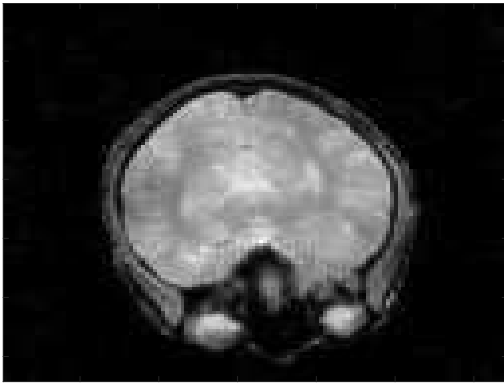
(b) Radial mask: 20.74%



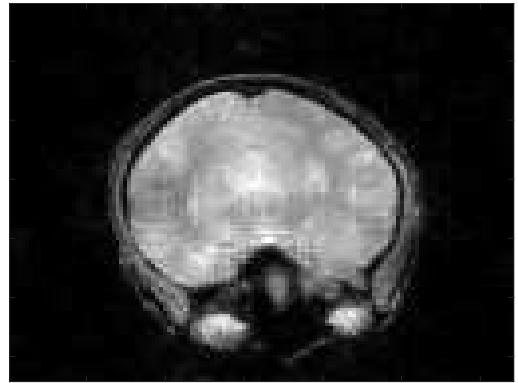
(c) Reconstruction: shearlets, spiral



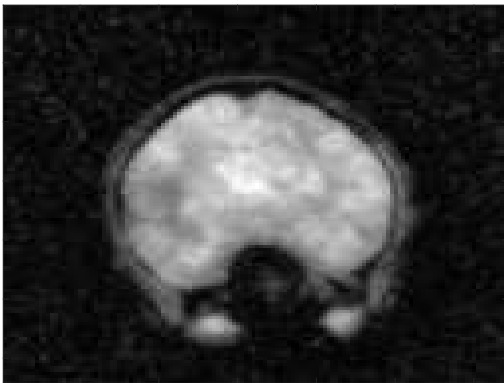
(d) Reconstruction: shearlets, radial



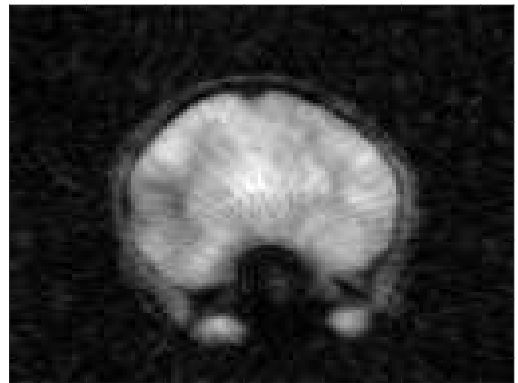
(e) Reconstruction: wavelets, spiral



(f) Reconstruction: wavelets, radial



(g) Reconstruction: Fourier inversion, spiral



(h) Reconstruction: Fourier inversion, radial

Figure 4: Reconstruction using a spiral and radial mask.

7 Future work

We like to address multiple extensions and future directions of the work presented in this paper.

7.1 Linear stable sampling rate

In this paper we have established a stable sampling rate for sampling with respect to complex exponentials and recovery using compactly supported shearlet frames. Our rate essentially differs from the rate for wavelets by an additional N^δ term and the lower frame bound. Although it seems plausible that the rate should depend on the stability of the redundant system, measured by the lower frame bound, it is interesting to see, whether it could be improved to actual linearity up to a dependency of the lower frame bound, i.e. erasing N^δ while keeping A_N . Furthermore, efficient numerical algorithms to compute the stable sampling rate for shearlets are currently missing. This issue is left for future work.

7.2 Nonuniform sampling

Moreover, we have restricted ourselves to the case of a uniform sampling pattern, in fact, the samples were assumed to be acquired on a grid. Although this is possible in practice, it is very unusual. Many MRI acquisition schemes are nowadays non-uniform schemes, e.g. spiral patterns or radial patterns are often used, see Figure 4a and Figure 4b. In [2, 1] the authors introduced *non-uniform generalized sampling* and proved stable reconstructions are possible when using *folded*, *periodic*, and *boundary* wavelets. We strongly believe that the methods used in the present paper can be used to obtain similar results for shearlets in the non-uniform case as well. This will be studied in an upcoming work.

7.3 Parabolic molecules

Shearlets are a special type of *parabolic molecules* [19]. It is therefore an interesting question whether our results can be extended to such systems. We explain briefly how such generalizations can be tackled.

First of all, the reconstruction space \mathcal{R}_N is spanned by the first N elements of the representation system. In our case of compactly supported shearlets it was built by considering all atoms whose support intersect a region of interest. If one considers atoms that are not compactly supported, then one has to make a careful choice of ordering the representation system and select the first N elements of the system. However, the argument for when a reconstruction from Fourier measurements is possible, is solely based on very good frequency localization of the atoms. Thus this argument is applicable to other systems such as wavelet frames or general parabolic molecules. However, for a qualitative judgement of the results one needs to relate the lower frame bound of the finite system to the approximation rate, similar as we did in (4.4).

7.4 Compressed sensing

Another perspective that is indispensable for reconstruction problems from Fourier measurements such as in MRI is *compressed sensing* cf. [10, 14]. Moreover, the new concepts of *asymptotic sparsity* and *asymptotic incoherence* introduced in [8] are perfectly suited for shearlets, in fact natural images are asymptotically sparse in shearlets and it can be shown that the cross gramian between complex exponentials and shearlets is asymptotically incoherent. Further investigations in this direction are also possible for future work.

As shown in the numerics, shearlets have a great potential in recovering real medical images from their subsampled Fourier data. The development of a sophisticated ℓ^1 algorithm is therefore of high importance. A mathematical analysis and a comparison of different recovery algorithms will be provided in an upcoming paper.

Acknowledgements

The author would like to thank Gitta Kutyniok for helpful discussions and the anonymous reviewers for critical comments that greatly improved the quality of the paper. Moreover, the author thanks Dr. Sina Straub from German Cancer Research Center (DKFZ) for providing the MRI data and Dr. Matthias Dieringer from Siemens AG, Healthcare Sector for providing a code to read the raw data into MATLAB. Further, the author acknowledges support from the Berlin Mathematical School as well as the DFG Collaborative Research Center TRR 109 "Discretization in Geometry and Dynamics".

References

- [1] B. Adcock, M. Gataric, and A. C. Hansen. Recovering piecewise smooth functions from nonuniform fourier measurement. *preprint*, 2014.
- [2] B. Adcock, M. Gataric, and A. C. Hansen. On stable reconstructions from univariate nonuniform fourier measurements. *SIAM Jour. Imag. Scienc.*, to appear.
- [3] B. Adcock and A. C. Hansen. A generalized sampling theorem for stable reconstructions in arbitrary bases. *J. Fourier Anal. Appl.*, 18(4):685–716, 2012.
- [4] B. Adcock and A. C. Hansen. Stable reconstructions in Hilbert spaces and the resolution of the Gibbs phenomenon. *Appl. Comput. Harmon. Anal.*, 32(3):357–388, 2012.
- [5] B. Adcock, A. C. Hansen, G. Kutyniok, and J. Ma. Linear stable sampling rate: Optimality of 2D wavelet reconstructions from fourier measurements. *SIAM J. Math. Anal.*, to appear.
- [6] B. Adcock, A. C. Hansen, and C. Poon. Beyond consistent reconstructions: optimality and sharp bounds for generalized sampling, and application to the uniform resampling problem. *SIAM J. Math. Anal.*, 45(5):3132–3167, 2013.
- [7] B. Adcock, A. C. Hansen, and C. Poon. On optimal wavelet reconstructions from fourier samples: Linearity and universality of the stable sampling rate. *Appl. Comput. Harmon. Anal.*, to appear.
- [8] B. Adcock, A. C. Hansen, C. Poon, and B. Roman. Breaking the coherence barrier: asymptotic incoherence and asymptotic sparsity in compressed sensing. *CoRR*, abs/1302.0561, 2013.
- [9] E. J. Candès and D. L. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 -singularities. *Comm. Pure Appl. Math*, pages 219–266, 2002.
- [10] E. J. Candès, J. K. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006.
- [11] O. Christensen. *An introduction to frames and Riesz bases*. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, Inc., Boston, MA, 2003.
- [12] I. Daubechies. *Ten lectures on wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
- [13] D. L. Donoho. Sparse components of images and optimal atomic decomposition. *Constr. Approx.*, 17(3):353–382, 2001.

- [14] D. L. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52:1289–1306, 2006.
- [15] Y. C. Eldar. Sampling with arbitrary sampling and reconstruction spaces and oblique dual frame vectors. *J. Fourier Anal. Appl.*, 9(1):77–96, 2003.
- [16] Y. C. Eldar and T. Werther. General framework for consistent sampling in Hilbert spaces. *Int. J. Wavelets Multiresolut. Inf. Process.*, 3(4):497–509, 2005.
- [17] L. Grafakos. *Classical Fourier analysis. 2nd ed.* Graduate Texts in Mathematics 249. New York, NY: Springer. xvis, 2008.
- [18] P. Grohs. Intrinsic localization of anisotropic frames. *Applied and Computational Harmonic Analysis*, 35(2):264 – 283, 2013.
- [19] P. Grohs and G. Kutyniok. Parabolic molecules. *Foundations of Computational Mathematics*, 14(2):299–337, 2014.
- [20] K. Guo, G. Kutyniok, and D. Labate. Sparse multidimensional representations using anisotropic dilation and shear operators. *Wavelets and splines: Athens 2005*, 1:189–201, 2006.
- [21] D. M. H. Jr. and J. B. Weaver. Two applications of wavelet transforms in magnetic resonance imaging. *IEEE Transactions on Information Theory*, 38(2):840–860, 1992.
- [22] P. Kittipoom, G. Kutyniok, and W.-Q. Lim. Construction of compactly supported shearlet frames. *Constructive Approximation*, 35(1):21–72, 2012.
- [23] G. Kutyniok and D. Labate. Introduction to shearlets. In *Shearlets*, Appl. Numer. Harmon. Anal., pages 1–38. Birkhäuser/Springer, New York, 2012.
- [24] G. Kutyniok and D. Labate. *Shearlets: Multiscale Analysis for Multivariate Data*. Birkhäuser Basel, 2012.
- [25] G. Kutyniok and W.-Q. Lim. Compactly supported shearlets are optimally sparse. *Journal of Approximation Theory*, 163(11):1564–1589, 2011.
- [26] D. Labate, W.-Q. Lim, G. Kutyniok, and G. Weiss. Sparse multidimensional representation using shearlets. *Wavelets XI., Proceedings of the SPIE*, 5914:254–262, 2005.
- [27] W.-Q. Lim. The discrete shearlet transform: a new directional transform and compactly supported shearlet frames. *IEEE Trans. Image Process.*, 19(5):1166–1180, 2010.
- [28] M. Lustig, D. Donoho, and J. Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn. Reson. Med.*, 58(6):1182–1195, 2007.
- [29] S. Mallat. *A wavelet tour of signal processing*. Elsevier/Academic Press, Amsterdam, third edition, 2009. The sparse way, With contributions from Gabriel Peyré.
- [30] L. Panych. Theoretical comparison of fourier and wavelet encoding in magnetic resonance imaging. *IEEE Trans Med Imaging.*, 15(2):141–53, 02 1996.
- [31] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger. Sense: Sensitivity encoding for fast mri. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 42(5), 1999.
- [32] W.-S. Tang. Oblique projections, biorthogonal Riesz bases and multiwavelets in Hilbert spaces. *Proc. Amer. Math. Soc.*, 128(2):463–473, 2000.

- [33] D. S. Taubman and M. W. Marcellin. *JPEG 2000: Image Compression Fundamentals, Standards and Practice*. Kluwer Academic Publishers, 2001.
- [34] M. Unser and A. Aldroubi. A general sampling theory for nonideal acquisition devices. *IEEE Transactions on Signal Processing*, 42(11):2915–2925, 1994.
- [35] Q. Wang, M. Zenge, H. Ertan, E. Mueller, and M. S. Nadar. Novel sampling strategies for sparse mr image reconstruction. *Proc. Intl. Soc. Mag. Reson. Med.* 22, 2014.